# Psychometrics validty and elaiability

Contents Introduction The selection of employees is one of the most significant tasks a human resources practitioner Is faced with. This affects the flow of employees entering and exiting the firm. Many Issues may arise If the recruiting process Is not In accordance with the South African Legislation, namely the Employment equity act and the labor relations act, which governs the reliability validity, bias and fairness of psychometric assessment measures.

These legislations have been brought about to protect employees againstdiscriminationand unfair practices which were experienced urine the previous dispensations. In accordance to the terms of the provisions of Section 8 of the Employment Equity Act (55 of 1998) " Psychological testing and other similar assessments of an employee are prohibited unless the test or assessment being used (a) has been scientifically shown to be valid and reliable; (b) can be applied fairly to all employees; and (c) is not biased against any employee or group".

The psychological measuring instrument we have chosen in accordance with HAPS is the PAIL B. The Ability Processing of Information and Learning Battery (PAIL-B) will e critically evaluated within this essay. This psychometric assessments battery can therefore assist recruiters who use these measures to identify employees who have the potential to grow and learn with In organizations.

Furthermore It should be noted that the PAIL B Is a cognitive measure and Is not only used for recruitment and selection in organizations, but can also be used for selection into schools, universities and other areas. Hence, this

essay will report the PAIL B through the following headings namely evaluating the PAIL B, Composition of the PAIL B, Validity, Reliability, Bias and Limitations.

Evaluating the PAIL B According to Foxtrot and Rood (201 3), It Is an assessment practitioner's duty to evaluate the Information offered about a measure and determine whether it Is valid and reliable for its intended purpose. Foxtrot and Rood (2013), further state that for evaluating a measure, some of the things that an assessment practitioner should consider are: how long ago it was developed; quality of manual contents; clarity of instructions and cultural appropriateness.

First conceptualized in 1994 by T. R Taylor, the PAIL B - Ability, Processing of Information and Learning Battery- (Taylor, n. . ), was designed as a set of tests with the purpose of assessing ones vital cognitive capableness. In order for the assessment to be most effective It should be administered on individuals with individuals with a minimum of twelve years educational background (Taylor, n. D. ).

The PAIL B is ideal for identifying those who are likely to master new cognitively challenging content in a training context and establishing levels in order to place people in the correct positions. Taylor (n. D. ), has identified three norms that the PAIL B makes use of, namely: stained (scale of 1 - 9) ; tens are used in the Flexibility-Accuracy-speed Tests (FAST), stained are used in the concept formation test; the memory test and Knowledge transfer test while percentiles are used in the curve of learning test.

According to Taylor (n. D. ), the PAIL B is divided into five test booklets and two ancillary booklets which make up eight scores namely: Abstract thinking; Speed of information processing; accuracy of information processing; cognitive flexibility; Performance gain in a learning task; final level of proficiency; Memory and understanding and Transfer of knowledge, which ill take approximately three hours and forty five minutes to administer.

Composition PAIL Battery Concept formation test This test was designed to assess one's ability to " think abstractly and conceptually: to form abstract concepts, reason hypothetically, theories, build scenarios (and) trace causes" (Taylor, p. 4, n. D. ). The test is comprised of thirty questions; each consisting of six depictions of similar nature the test taker must identify the depiction that does not share a characteristic that the rest of the depictions share (Taylor, n. D. ). Flexibility-Accuracy-speed Tests (FAST) Taylor (n. D. Suggests that " this battery within a battery measures speed (quickness) and accuracy of information processing, and cognitive flexibility'. The FAST test is made up of four individual assessments namely: Series; Mirror image; Transformations and combined tests. All four assessments are time sensitive and have been designed in such a way that it is very rare for a test taker to actually complete the entire assessment. It uses shapes of different sizes which may contain either a dot or line in the center. The basic idea of the tests is to identify a pattern and find the omitted depiction.

Curve of learning According to Taylor (n. D. ), this test focuses on a learning potential, it aims to assess ones capacity at which they are able to master new skills. It looks at future achievement potential rather than the abilities

that the person already has. The test is split into four timed sessions which requires the test taker to decode a series of paired images into another set of images and once again decode these images to a set of words. Images are decoded with the aid of the first ancillary booklet, the dictionary.

Memory test Directly after the test taker has completed the curve of learning test, the memory test s administered. It follows the same concept as the curve of learning where the test takers are required to decode images to words; however the dictionary is now taken away. The performance of the test taker on this test reflects the extent to which the test taker has understood the logical relation between the symbols and words. Knowledge transfer test According to Ferguson (1956, as cited in Taylor, n. D. ), transferring knowledge and skills to similar areas or situations is a vital process of cognitive development.

The knowledge transfer test, as the name suggests, measures this ability. The test insists of a series of connected depictions referred to as " pieces of equipment" (Taylor, p. 19, n. D. ), which have a specific feature in addition to a basic shape. The test taker is required to categorize them under symbols. Test takers are also given the second ancillary booklet. Validity established to ensure that the test is valid for the purpose it is to be used for. Foxtrot and Rood (2013) state that the " the validity of a measure concerns what the test measures and how well it does so".

In the studies consulted it has been evident that construct and criterion validity were shown to be present in the PAIL B assessment. The construct validity of a measure is the extent to which it measures the theoretical

construct or trait that it is supposed to measure (Foxtrot & Rood 2013). The second validation measure of criterion validity was defined by Paella and Wren (2005) who stated that " Criterion-Related Validity is used to predict future or current performance". The method that used to determine criterion related validity is predictive validity.

Murphy and Adversaries (2005) define predictive validity as a method of determining criterion validity. It also used to determine the correlation of test takers test score and there criterion related scores. Taylor (1995) investigated the validity on the SOFT, where he gave the measure to 33 first-year university students who had been accepted into the university on merits other than their grade twelve results. Taylor correlated the marks from their SOFT assessments and the marks of the course they took; which were to improve their logical thinking and reasoning skills. Therefore the correlation was 0. 4 (p = 0. 012). Taylor (1995, as cited in Astrakhan, 1998) found in another study which investigated the validity on the Curve of Learning and Memory and Understanding tests was inducted using a sample of 110 workers from a beverage manufacturing firm. The criteria for evaluating workers included facets such as their capacity to learn new procedures and concepts, to understand why things happen in the firm as a whole, and their capacity to plan and organism. These results averaged correlations of 0. 35. The low correlation can be attributed to the fact that a diverse sample was not used.

A further study by Taylor (1995) found criterion scores which was given to 43 employees who were enrolled in a course designed to prepare them for a

promotion in Junior management positions. The correlations here were reported to be 0. 67 and 0. 79 respectively, which can be interrupted to prove to be an recreate predictor of performance. In an additional study conducted by Lopes, Rood and Maier (2001) on the predictive validity of the PAIL-B in a financial institution; the purpose was to assess the predictive validity of the PAIL test battery, in order to identify learning potential.

A sample of 235 successful Job applicants were used to complete the test battery and found the predictive validity of the test battery was assessed using a canonical discriminate analysis procedure. The procedure was adopted in view of the nominal strength of the manager's ratings, and due to the limited sample size the 5 point rating scale was eventually collapsed too 2 point classification. Reliability It should be noted that an assessment is reliable if it measures the same construct in a consistent and precise manner over time.

Foxtrot and Rood (2009) define reliability of a measure as " the consistency to which it measures whatever it measures". Split - half reliability was a major psychometric property of reliability used among majority of the literature we consulted. In the PAIL B, (Taylor 1995) elucidates that split half reliability was used to investigate whether or not the PAIL-B is reliable. Foxtrot and Rood (2013, p. 47) define split-half reliability as " obtained by splitting the measure into two equivalents (after a single administration of the test) and computing the correlation coefficient between these two sets of scores".

During Taylor' investigations into the reliability of the PAIL B, he used a sample of six groups to test reliability coefficient of the flexibility, accuracy and speed test and the knowledge transfer test. These have reliability coefficients room a low of 0. 70 - too high of 0. 86 and 0. 71 - 0. 84 respectively Taylor (1995). In a study done by the defense force which lasted over a period of three years with new recruits. The purpose was to determine whether the psychometric evaluation processes can reliably predict the learning potential of first year recruits at the academy.

The FAST considered the following; firstly, the PAIL B investigated whether the FAST has a positive effect on how quickly recruits learn new abilities. It was found that a significant relationship with a reliability coefficient of (r= O, 491) exists between legibility of information processing and steepness of the learning curve. This therefore is below the accepted reliability coefficient of Secondly, it was found that a strong relationship with a reliability coefficient of is apparent between speed of information processing and the total amount of work completed by the recruits.

Lastly, it was determined that the small relationship with a reliability coefficient of exists between accuracy of information processing and steepness of the learning curve. This therefore is below the accepted reliability coefficient of However, the results concluded that three components of the FAST, are accurate in predicting how quickly new recruits in the defense force will develop new competencies. The findings also further indicated that the accuracy with which information is processed has a

minimal influence on the rate a recruit will develop new competencies (Preterits 2010).

In terms of the knowledge transfer test which investigated if there was a transfer of knowledge to crystallized abilities. Meaning it investigated if there was a transfer of what the recruits learnt and how they apply it in combat situations. Preterits (2010) defines crystallized abilities as " are peccadillo insight or understanding and knowledge that emerge via transfer from existing knowledge and that is subsequently, successfully stored in memory'. The Memory and Understanding sub-test of the PAIL-B was used to measure crystallized ability of recruits.

It was found that a positive relationship exists between the transferring of knowledge in what the recruit learnt and crystallized abilities. The reliability coefficient was reported as a positive directional relationship between transfer of knowledge and crystallized abilities. A substantial relationship with a liability coefficient of 5) exists between memory and understanding and crystallized abilities. This therefore suggests that a moderate correlation exists.

In terms of the curve of learning, it was found that prior learning has a positive directional effect on learning performance thus the results indicate a substantial relationship and moderate correlation with a reliability coefficient of In concluding with this study, it can be said that the defense force's use of the PAIL B (Preterits 2010). A De Geode and Thereon (2010) study concurred with Preterits (2010) where a non- arability sample of 434 new recruits from

the South African Police Service Training College in Philipp, Cape Town was used.

Even though the size of the selected sample is quite acceptable, making use of a non-probability sampling of the target population, caution should be taken when making generalizations. De Geode and Thereon (2010), found that a score of reliability score of (r= 0. 45). This suggests that a question mark hangs over the success with which at least some of the concealed variables comprising the results of the learning potential police recruits. Standard Error of Measurement Foxtrot and Rood (2013, p. 49), " explain that the standard error of measurement indicates the band of error around each obtained score, and examiners should aware of the standard error of measurement for each subtest before interpreting the test - takers score". Therefore, assessors must be cognizant of the test takers history and current circumstances. Factors such asculture, transient conditions, prior learning and test wishes can have an impact on the variance between the true score (obtained under perfect conditions) and the obtained score.

Preterits (2010) outlines hat prior learning of an individual and their familiarity with taking assessment has a significant impact on their ability to perform in test conditions. While Dodos (2000) was of the view that a testes culture as well as environmental factors will also affect the scores of the historically disadvantaged people of South Africa. Bias Prior (2011) explains that bias " implies that test scores obtained for various subgroups of a given population cannot be interpreted in the same way across the groups".

Taylor (1995) suggests PAIL-B was designed as a learning potential test and therefore emits any biased based on cultural differences. This is a result of the test being a non-verbal test, except for the instructions, and the test comprises of mainly geometric depictions thus language does not become an issue of concern. Astrakhan (2008) concurs with Taylor (1995) in a study conducted with a sample of 400 individuals, 66 testes had African surnames while the large majority can be classified as white.

The data analysis for both race groups were highly correlated indicating that there is no potential for bias. However, it should be kept in mind that his was not a representative sample. Further studies were consulted to investigate the potential bias in the PAIL B. A sample of 20 psychological professionals from various fields, were asked to investigate the cultural bias of the PAIL-B; found that 6 out of the 20 felt that the test was bias (Dodos 2000). Thus, it can be stated that there is a potential for bias based on one's culture.

Similarly, Preterits (2010) concurs with Dodos as he found that the PAIL B was accused of being biased and under representing the cognitive capacity of individuals from historically disadvantaged backgrounds. Thus, in order to bring recruitment practices in line with legislation in the Employment Equity act, these tests was subsequently replaced with a selection battery thought to be less susceptible to culture, race and gender bias. This resulted in the measure being removed from use in the context of the defense force.

Limitations of the PAIL B diverse enough for the representative target population. This is further verified by Astrakhan (2008), who also did not make use of a diverse sample. Therefore based on the literature from these

authors, it is evident that accurate conclusions cannot be ran indicating that there are limitations in the above studies. Conclusion In the end the results show that, the PAIL-B is able to predict the performance of individuals not only in certain institutions but for any selection at an accurate level and therefore makes the battery a vital instrument to use.

It is evident that the PAIL- B is a somewhat outdated measure but still proves to be valid and reliable in measuring cognitive abilities today. However, caution could be taken when administering the PAIL-B, as some authors have found that bias is present in historically disadvantaged groups. This essay therefore reported on the PAIL B through evaluation of the PAIL B, Composition of the PAIL B, Validity, Reliability, Bias and Limitations. Recommendations Firstly, it should be noted that the PAIL B is an outdated selection battery.

In order for organizations to make fair decisions in line with the Employment Equity Act, a more relevant battery needs to be considered. Secondly, it should also not be used in its individual capacity within the recruitment and selection process and is it advisable to be used in harmony with other valid information such as candidates' curriculum it's and other test results. Thirdly, the use of the PAIL can be considered bias in instances where people from different cultures and race groups are affected.

In addition, Astrakhan (2008), De Geode and Thereon (2010), should make use of a more representative sample in order to draw conclusions about the reliability of their studies. Lastly, we also propose that measures within the battery not require such strict prior learning criteria as these have been

shown to bias the historically disadvantaged individuals who have not had exposure to prior learning. Reference List psychometric test administrators toward the PAIL B as a culture fair assessment with special reference to the employment equity act.